

Katie Vicari & Erin Dewalt/Nature Publishing Group

The methods for assessing methylation include enzymatic and chemical enrichment methods, DNA immunoprecipitation, quantitative PCR, hybridization, microarrays and sequencing.

A menu of profiling choices

Among the many methods to assess methylation, there are, for example, enzymatic and chemical enrichment methods. Using methylated DNA immunoprecipitation lets scientists draw DNA fragments out of solution with the help of antibodies against specific modifications. Further analysis can involve techniques as varied as quantitative PCR (qPCR), hybridization, microarrays or sequencing.

Many assays rely on bisulfite treatment. Bisulfite conversion helps to reveal the ratio of methylated and unmethylated cytosines in a genetic region: DNA is sheared or sonicated into pieces and then treated. Bisulfite leaves 5mC unaltered, while unmethylated cytosine is changed to uracil and then to thymine in an amplification step with primers.

Microarrays can be used to test the methylation status of thousands of DNA fragments at once after bisulfite treatment. The available platforms include the Agilent CpG island arrays and methylation arrays with CpG island and promoter regions and Infinium bead methylation arrays from Illumina.

As Feinberg explains, he and his team had been using arrays, but they have switched almost entirely to bisulfite sequencing now that its price has come down. Bisulfite conversion followed by sequencing has been in use since it was developed in 1992 (ref. 2), and it remains the gold standard for flagging methylated cytosines. “If it survives that long and is still

called the gold-standard method, it must be doing something right,” says Reik.

A plethora of methods leverage bisulfite treatment followed by sequencing, and they help to map methylation sites on a genome-wide basis or in a more locus-specific fashion. One limitation that held the method back was that it did not distinguish 5mC and 5hmC. That issue was overcome in 2012, when scientists at the University of Cambridge and the Babraham Institute developed oxidative bisulfite sequencing (oxBS-seq), in which 5hmC is first converted to 5fC before bisulfite treatment³.

Bisulfite conversion has a fan base, but there are technical challenges. For example, says Reik, researchers will not always know if the conversion has been completed and all nonmethylated cytosines have been converted to uracils. “That’s a very important quality measurement,” he says. Also, bisulfite conversion is harsh, and plenty of DNA is lost in the process. The harsher the treatment, the better the conversion will likely be but the worse DNA loss becomes: “It’s a balance,” he says.

The process also has biases, says Reik. In mammalian cells, CpG islands, with high CG density, are found all across the genome, but regions outside these islands, for example in introns, have a much lower CG density. This variability means that the method can represent one genomic region better than another. “That’s standing there in the room at the moment,” says Reik. “People don’t talk about it much because they don’t have a solution to it, basically.” He advises bearing this in mind when analyzing data. There are computational workarounds, but compensating for a problem is never ideal, he says.

Bisulfite-free approaches

A number of bisulfite-free methods for DNA methylation profiling have also emerged. For example, Chuan He of the University of Chicago and his colleagues at several institutions in China developed fC-CET, a bisulfite-free approach that detects 5fC. They label 5fC with a chemical that they synthesized, an azido derivative of 1,3-indandione, to convert 5fC without DNA degradation. The scientists can then use sequencing to map the 5fC locations.

The desire to profile DNA methylation is growing in cancer research and in developmental biology, where scientists often have only small tissue samples for their assays, says Sriharsa Pradhan, who runs epigenetics

research at New England Biolabs (NEB). Generally, he hopes bisulfite-free DNA methylation assays will help scientists figure out how modifications such as methylation are maintained and inherited.

When researchers have more than a few micrograms of material, bisulfite conversion might be appropriate. Bisulfite is also a cheap chemical. But, says Pradhan, for work on single cells or in situations with tiny amounts of sample, it may be preferable to use enzymes. Enzymatic assays allow scientists to use longer DNA sequence fragments and to perform DNA phasing, in which maternal and paternal alleles can be distinguished, he says.

NEB offers a number of restriction enzymes that cleave methylated DNA at specific locations. Scientists might draw on the MspJ1 or the PvuRts1I family of restriction enzymes to get a locus-specific single-base resolution readout of a methylated site. These enzymes will cut the DNA irrespective of sequence, wherever they find a 5mC, 5hmC or 5fC, says Pradhan. Researchers can then use qPCR to quantify methylation or sequencing to map it. The enzymatic approach helps provide a quantitative readout at a particular site where there might be low abundance of methylation in either a CpG island or regions of no CpG context, he says. Labs can use enzymes to quantify and assess methylation changes as a cell population reacts to a perturbation. An NEB team has shown how to use enzymes to map 5hmC and 5fC genome-wide at single-base resolution⁴.

Drawing on crystal structures to see how the enzymes contact the DNA, Pradhan and colleagues have been studying the enzymes’ behavior. “When you are going for a single site, you have to be sure the enzyme behaves in the right way,” he says.

Used incorrectly, the enzyme risks cutting an unmethylated site, or it might bind to the methylated site but not cut, says Pradhan. The NEB researchers have brought down the false positive rate to 2%, he says, which hints at potential clinical applications as well as the possibility of multiplexed experiments to cut and assess many methylated sites in parallel. As controls for a qPCR assay, the company offers DNA sequence with known methylation levels.

With sequence analysis, it is often not easy to deconvolute the signal at methylated sites, says Pradhan. Cytosines close to one another may be methylated but have unmethylated cytosines interspersed

between them. But, he says, computational methods will emerge to address these issues at a given locus.

Bisulfite-free sequencing

To map DNA methylation sites without bisulfite conversion, some labs turn to long-read sequencing with the Pacific Biosciences (PacBio) platform. In contrast to sequencing technologies in which the amplification step in sample prep removes methylation patterns, the company's Single-Molecule, Real-Time (SMRT) sequencing lets researchers tally methylation profiles as sequencing progresses by tracking the incorporation of fluorescently labeled nucleotides. "The direct detection of methylation with SMRT sequencing gives you access to many more modifications than just 5mC," says Jonas Korlach, chief scientific officer of PacBio.

As part of investigating a particularly virulent *Escherichia coli* outbreak, Eric Schadt, Korlach's predecessor, who is now at the Icahn School of Medicine at New York's Mount Sinai Hospital, used the platform to investigate why this particular *E. coli* strain was virulent enough to kill 50 people. Sequencing showed that the bacterium contains an insertion from a prophage element that makes it enterohemorrhagic. And there is a unique identifier, a signature specific to this strain's methylation pattern, its N⁶-methyladenine levels⁵.

Since then, a team at the University of Leicester and several other institutions has used SMRT sequencing to study the pneumonia-causing bacterium *Streptococcus pneumoniae* and found distinctive methylation patterns that correspond to levels of its virulence. PacBio has also worked with a group at Griffith University in Australia and Ohio State University and found how variable N⁶-adenine methylation patterns change gene expression in *Neisseria meningitidis*, the bacterium that causes meningitis. The variation contributes to antibiotic resistance. As Korlach explains, these variations—phasevariations—regulate gene expression and are the mode through which host adaptation and virulence are regulated in many human pathogens.

Bisulfite conversion combined with SMRT sequencing lets labs sequence longer DNA fragments than short-read technologies, and they can perform epigenetic phasing, says Korlach—separately assessing the methylation patterns for the maternal and paternal alleles.

Another bisulfite-free sequencing approach involves direct readout from nanopore sequencing. With this technique, methylation and modification data come along "for free" as part of regular normal genome sequencing, says Johns Hopkins researcher Timp. In his current work, he is using nanopore sequencing to characterize how DNA modifications such as methylation change the structure of DNA and hence alter the electrical signal output of the sequencer.

Timp and others have found that methylated DNA goes through the nanopore more readily than unmethylated DNA, and a number of labs have noted that methylation appears to affect the stiffness of DNA molecules, he says. He has used sequencing data obtained with an Illumina platform to check the nanopore sequencing and methylation profiling data results on Oxford Nanopore's MinION. He sees high concordance in the tests thus far. Potentially, he says, the approach could detect modifications to RNA, too.

Profile patterns

Cancer cells are unlike healthy cells in terms of their genes, their epigenetic modifications and their appearance. Early in the twentieth century and long before the advent of sequencing, biologist Theodor Boveri noted the different appearance of the nucleus in cancer cells. Sequencing has subsequently shown the genetic disarray in cancer cells.

Mutations turn a cell cancerous, but in Feinberg's view, epigenetic modifications will play a large role in deciphering many remaining cancer biology questions. He and his team have found long stretches of hypomethylation in cancer genomes, which appear to activate certain oncogenes. There is also much methylation variability in loci directly adjacent to CpG islands.

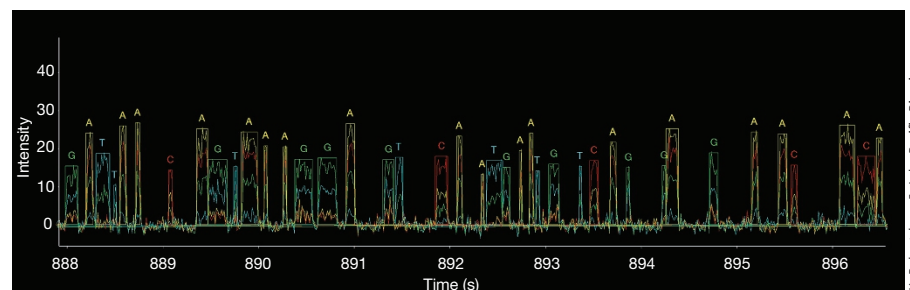
As Feinberg and his team explore these patterns, they have been hunting for the

information's deeper structure. On a visit to Darwin's and Newton's graves at Westminster Abbey, the idea came together for Feinberg. "Maybe there is a selective advantage for some genes to have epigenetic variability," he says. Epigenetic plasticity could be an important trait that enables tumor propagation⁶.

Increased epigenetic variability at a specific gene locus is also a possible predictor of cancer susceptibility. This idea stems from the finding that there is increased variability of gene expression in cancer cells at loci where DNA methylation in healthy cells varies. These loci of increased methylation variability "are there not to give you cancer," says Feinberg. These regions are called up as part of the biological response to stress or injury. Repeated insults might increase the derailment risk at these loci.

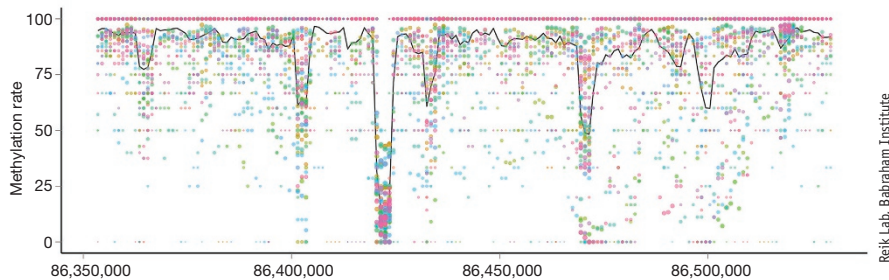
Typically labs tally the ratio of methylated to unmethylated sites, which is not that informative, says Feinberg. At an individual locus, methylation is not a quantity—a cytosine molecule is not in some cases a little methylated and very methylated in others. "It's either methylated or it isn't," he says. The more relevant quantitative information, Feinberg says, is coded in the pattern of methylation states, and he hopes to discover how this pattern in a cell population affects cell behavior and tissue response.

Along with his Hopkins colleague electrical engineer John Goutsias, Feinberg is developing mathematical methods as well as assays to deconvolute the methylation signals at genetic loci and across many cells. As he deciphers what drives the probability of DNA methylation in a cell population, he hopes to learn about the nature of epigenetic information that is transmitted during cell division from an information-theory point of view. "Part of that will be single-cell analysis; obviously that's going to be very helpful for this kind of study," he says. Single-cell approaches can help to check predictions, for example.



By tracking the kinetic signatures that identify methylated bases during sequencing, Single-Molecule, Real-Time (SMRT) sequencing helps scientists find modified bases.

K. Robertshaw, J. Lyle, Pacific Biosciences



Reik Lab, Babraham Institute

A full picture of the role of epigenetic marks could emerge from capturing the states of single cells in terms of their transcriptome and methylome. Here, each colored dot represents the methylation status of a single cell at a specific locus. The black curve represents mean methylation across all cells.

Toward single-cell analysis

Daniel Messerschmidt, a researcher at the Institute of Molecular and Cell Biology at A*Star Singapore, and his team recognized that methylation can vary between individual cells and developed an assay to probe the methylation of individual loci on this single-cell level in mouse embryos⁷. The team explored patterns involved in imprinting, in which one inherited allele at a given locus is silenced through methylation. Disease and disorders can result when these patterns are disrupted. The scientists used single-cell DNA digestion with a methylation-sensitive enzyme, followed by qPCR amplification, to assess multiple loci of interest in a microfluidic device.

Functional studies of DNA methylation are enmeshed with the challenge of understanding the role of base modifications. Methylation can be part of the differentiation process, which can give cues for changes in cell state, says Reik. Those shifts might be transient as a result of quick turnover, but some modifications might not be as transient as scientists have assumed. For example, 5fC is involved in DNA demethylation and also has a role as a more stable modification. Although the actual function of 5fC remains puzzling, Reik believes it is more than a chemical intermediate in the methylation-demethylation cycle.

In the team's quest to quantify the rare modifications 5fC and 5caC, Shankar Balasubramanian at the University of Cambridge, along with Reik's team and scientists at two other UK institutions, fed mice an isotope-labeled amino acid, [*methy*-¹³C₃]-L-methionine, and then tracked with mass spectrometry the changes in

methylated cytosine in mouse embryonic tissue and in the animals through adulthood to 15 weeks of age. Whereas 5mC, 5hmC and 5fC can be identified in tissues such as kidney and colon in mouse pups, 5fC is always less abundant than 5mC or 5hmC. But in adult brain, 5fC is the most abundant modification. This variation is informative, says Reik, and it may be what he calls context dependent. Methylation at a particular locus can be a cue for developmental phases in certain organs. In the brain, where there is little cell division in adult animals, the function of this methylation remains to be uncovered.

"If you want to know how many cytosines in this genome are methylated very quantitatively, then mass spec is a fabulous method for that," says Reik, given that each modified base has a different mass. The approach delivers an average methylation pattern across the tissue. On the flip side, he says, that means that "you're ignoring the heterogeneity that exists."

Single-cell methods, of which there are a few, are a way to close in on this heterogeneity. That view can help reveal the details of the cell fate decisions that turn one cell into a heart cell and another into a liver cell. "We're beginning to see that, we think, so we're hugely excited about that," says Reik. The team has advanced a single-cell approach with a profile of 61 mouse embryonic stem cells. They captured the cells' distinct cell states in terms of their transcriptome and methylome⁸.

Their method, scM&T-seq, lets them connect methylation data from cells with transcriptional output from those same cells. It's an integrated analysis and an alternative to the practice of averaging

methylation patterns across many cells, says Reik.

The method involves single-cell bisulfite sequencing along with a protocol the team previously developed called genome and transcriptome sequencing (G&T-seq) to first physically separate and amplify DNA of single cells. The process separates a single cell's polyadenylated (poly(A)) RNA from its genomic DNA using a biotinylated oligo-dT primer. Then the genome and the transcriptome are amplified in parallel and sequenced.

Speaking generally about single-cell methods, UCL's Beck says that obtaining genomes, transcriptomes and methylomes from the same cell is of great interest in the community of researchers who work on blood-based assays called liquid biopsies. Single-cell methylation profiling matters, he says, because "we need a lot of improvement here to do what many want which is to use liquid biopsies to analyze essentially all human cell types."

Labs studying more general changes on a whole-organ level, such as methylation differences between disease and healthy states, might not need this single-cell approach, says Reik. But he hopes that labs pursuing biological questions in which epigenetic heterogeneity matters, such as aging, development and cancer, will find it a powerful way to link methylation patterns and transcription levels in single cells.

Ultimately, methylation profiles of single cells will help researchers understand methylation at specific loci, which can, in turn, expand their view of the role of the methylome more broadly in cell populations and tissue.

1. Plongthongkum, N., Diep, D.H. & Zhang, K. *Nat. Rev. Genet.* **15**, 647–661 (2014).
2. Booth, M.J. *et al. Science* **336**, 934–937 (2012).
3. Frommer, M. *et al. Proc. Natl. Acad. Sci. USA* **89**, 1827–1831 (1992).
4. Sun, Z. *et al. Mol. Cell* **57**, 750–761 (2015).
5. Schadt, E.E. *et al. Genome Res.* **23**, 129–141 (2013).
6. Feinberg, A.P. & Irizarry, R.A. *Proc. Natl. Acad. Sci. USA* **107** (Suppl. 1), 1757–1764 (2010).
7. Lorthongpanich, C. *et al. Science* **341**, 1110–1112 (2013).
8. Angermueller, C. *et al. Nat. Methods* doi:10.1038/nmeth.3728 (11 January 2016).

Vivien Marx is technology editor for *Nature Methods* (v.marx@us.nature.com).